AI TECHNIQUES FOR SOFTWARE COST ESTIMATION

Dr. Simon Wright
simon@symtech.ltd.uk
ARTIFICIAL INTELLIGENCE HAS BECOME UBQUITOUS

Every time we do a search in the background AI services swing into action. So, could AI help with cost estimation?
CAN ALEXA ESTIMATE COST?
COST ESTIMATION PROCESS

Capture the User need
- User Stories

Size the User Stories
- Function Points

Use benchmarks to estimate
- Quality
- Time
- Cost
THE ATTRIBUTES OF AN ESTIMATE

**Precise**

Precision = \( f(\text{completeness}) \)

- If a user story is missed, then the estimate is imprecise

**Accurate**

Accuracy = \( f(\text{correctness}) \)

- If a user story is incorrectly expressed, then the estimate is inaccurate

**Repeatable**

Repeatability = \( f(\text{consistency}) \)

- If a user story uses terms inconsistently then different estimators are likely to use different interpretations and thereby produce different estimates.
THE AI TECHNIQUE OF NATURAL LANGUAGE PROCESSING

• A **Correct** user story

• A **Complete** sets of user stories

• A **Consistent** set of user stories
WE ALL CAN SPOT “POOR” STATEMENTS OF USER NEED

1. It will implement the same set of functions as its predecessor
2. Relevant sets of project data will be sorted as quickly as possible
3. Appropriate standards shall be used when necessary
4. Careful consideration should be given to using a high-bandwidth channel
5. Windows 10 shall be installed on all PCs

Poorly phrased User Needs contain **words, constructs or phrases** that must be **avoided**.
WE ALL CAN SPOT “POOR” STATEMENTS OF USER NEED

1. It will implement the same set of functions as its predecessor

2. Relevant sets of project data will be sorted as quickly as possible

3. Appropriate standards shall be used when necessary

4. Careful consideration should be given to using a high-bandwidth channel

5. Windows 10 shall be installed on all PCs

• 5 words to change

• 6 words to change

• 3 words to change

• 4 words to change

• 1 word to change.
TEXT QUALITY METRIC

The number of times a word or phrase occurs in the “avoid” list

• Quality Level **Low** = Quality Score more than or equal to 5
• Quality Level **Medium** = Quality Score more than 1 and less than 5
• Quality Level **High** = Quality Score more than 0 and less than 1

<table>
<thead>
<tr>
<th>Score</th>
<th>Q. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Low</td>
</tr>
<tr>
<td>6</td>
<td>Low</td>
</tr>
<tr>
<td>3</td>
<td>Medium</td>
</tr>
<tr>
<td>4</td>
<td>Medium</td>
</tr>
<tr>
<td>1</td>
<td>Medium</td>
</tr>
<tr>
<td><strong>Total 19</strong></td>
<td><strong>Av. 3.8</strong></td>
</tr>
</tbody>
</table>

1. It will implement the same set of functions as its predecessor
2. Relevant sets of project data will be sorted as quickly as possible
3. Appropriate standards shall be used when necessary
4. Careful consideration should be given to using a high-bandwidth channel
5. Windows 10 shall be installed on all PCs

The next step is syntax.
THE SYNTAX OF A USER STORY

• As an author of user stories I want to maintain high quality User Stories

• As a Product Owner I need to list User Stories alphabetically
## THE SYNTAX OF A USER STORY

<table>
<thead>
<tr>
<th>THE ACTOR</th>
<th>IMPERATIVE</th>
<th>AN ACTION</th>
<th>AN OBJECT (+CONDITION)</th>
</tr>
</thead>
<tbody>
<tr>
<td>As a US Author</td>
<td>I want to</td>
<td>maintain</td>
<td>User Stories</td>
</tr>
<tr>
<td>As a Product Owner</td>
<td>I need to</td>
<td>list</td>
<td>User Stories alphabetically</td>
</tr>
</tbody>
</table>

\(<\text{ACTOR}>\) \(<\text{ACTION}>\) \(<\text{OBJECT}>\) and \\
\(<\text{TARGET}>\) \(<\text{PERFORMANCE}>\) \(<\text{TRIGGER}>\) \(<\text{CONSTRAINT}>\) \(<\text{RATIONALE}>\)*

SYNTAX FOR EFFECTIVE USER STORIES

• As a [Actor – who/what does the action]
• I want to [Action – what happens e.g. store, update, search]
• a [Direct Object – what is acted upon]
• in/on/with [Indirect Object – what is acted upon]
• on the [Target – where the output is sent]
• with [Performance - frequency and/or quality characteristic]
• when [Trigger – causes of action]
• unless / even if [Constraint – business rule or limiting factor]
• So that [Rationale - description of value or benefit is achieved]

A uniform syntax increases clarity, consistency and completeness.
EXAMPLE USER STORY

As a **library_user** I want to **search** for **books** in the library catalogue on the **computer screen** with speed and ease-of-use when I provide a book title even if **the book title is incomplete** so that I can find all books with similar titles

<ACTOR><ACTION><OBJECT(s)> //
<TARGET><PERFORMANCE><TRIGGER><CONSTRAINT><RATIONALE>
EXAMPLE USE CASE

I provide a book title and I begin a search for books in the library catalogue on the computer screen even if the book title is incomplete so that I can find all books with similar titles.

<TRIGGER><ACTOR><ACTION><OBJECT><TARGET> <CONSTRAINT> <RATIONALE>
As a library_user I want to search for books in the library catalogue on the computer screen with speed and ease-of-use when I provide a book title even if the book title is incomplete so that I can find all books with similar titles

Who & What
When & How
Why

The next step is semantics.
SEMANTIC ANALYSIS OF 3 WORDS

- The three words ‘book’, ‘library’ and ‘a’ can be ordered in six ways.
- Only three are syntactically correct.

- “A library book” (An object)
- “A book library” (A descriptor)
- “Book a library” (An action)
### Correctness and Consistency

#### Book the book from the book library

<table>
<thead>
<tr>
<th>Phrase</th>
<th>Syntactically Correct?</th>
<th>Semantically Consistent?</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>“A library book”</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>“A book library”</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>“Book a library”</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>“Library a book”</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>“Library book a”</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>“book Library a”</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Book the book from the book library</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Reserve the book from the document library</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
THE STRATEGY

• NLP to identify Users in the User Stories
• NLP to identify Processes in the User Stories
• NLP to identify Entities in the User Stories

• Correctness analysis to identify incorrectly expressed user stories
• Completeness analysis to identify incomplete epics
• Consistency analysis to identify inconsistent terms in an epic

• Allowing Functional size and Cost Estimates to be made.
IDENTIFYING USERS

• Users come in many shapes and sizes
  • Normal Operator
  • Maintenance Operator
  • Operational Support
  • Interfacing System / API
  • Clock
  • Other Software etc.
IDENTIFYING USERS

Subject
• As a library user I want to search for books
• on the computer screen
• with speed and ease-of-use
• when I provide a book title
• even if the book title is incomplete
• so that I can find all books with similar titles

Who & What

When & How

Why
IDENTIFYING PROCESSES

User(s)
Humans
Other Systems
Other Software

DATA → PROCESS → DATA

STORAGE

Data / Transactional / Functional
IDENTIFYING PROCESSES

SEARCH TERM → Search → MATCHING DATA → Catalogue

data

SEARCH TERM

MATCHING DATA
IDENTIFYING PROCESSES

Verb
• As a **library user** I want to **search** for **books**

Who & What
• on the computer screen
• with speed and ease-of-use
• when I provide a book title
• even if the book title is incomplete
• so that I can find all books with similar titles

When & How

Why
IDENTIFYING ENTITIES

Search

BOOK SEARCH TERM

MATCHING BOOK DATA

BOOK DATA

BOOK Catalogue
As a library user I want to search for books

- on the computer screen
- with speed and ease-of-use
- when I provide a book title
- even if the book title is incomplete
- so that I can find all books with similar titles
THE STRATEGY

- NLP to identify Users in the User Stories
- NLP to identify Processes in the User Stories
- NLP to identify Entities in the User Stories

- Correctness analysis to identify incorrectly expressed user stories
- Completeness analysis to identify incomplete epics
- Consistency analysis to identify inconsistent terms in an epic

- Allowing Functional size and Cost Estimates to be made.
THE TACTICS

• NLP to identify **Subjects** in the User Stories
• NLP to identify **Verbs** in the User Stories
• NLP to identify **Objects** in the User Stories

• Correctness analysis to identify incorrectly expressed user stories
• Completeness analysis to identify incomplete epics
• Consistency analysis to identify inconsistent terms in an epic

• Allowing Functional size and Cost Estimates to be made.
THE SOLUTION

• NLP to identify **Subjects** in the User Stories
• NLP to identify **Verbs** in the User Stories
• NLP to identify **Objects** in the User Stories

• **Correctness** analysis to identify incorrectly expressed user stories
• **Completeness** analysis to identify incomplete epics
• **Consistency** analysis to identify inconsistent terms in an epic

• Allowing Functional size and Cost Estimates to be made.
Correctness

• Has every User Story at least
  • one User, (subject)?
  • one Process, (verb)? and
  • one Entity, (object)?

Consistency

• Have Users been named consistently?
• Have Processes been named consistently?
• Have Entities been named consistently?

Completeness

• Has every Users been the subject of a user story?
• Does every Entity have a “create” verb in a user story?
• Does every Entity have a “read” verb in a user story?
• Does every Entity have a “update” verb in a user story?
• Does every Entity have a “delete” verb in a user story?
As a library user I want to search for books on the computer screen with speed and ease-of-use when I provide a book title even if the book title is incomplete so that I can find all books with similar titles.
Data Movements found by NLP
Functional Requirement:

As a library user I want to search for a book

Notes, scenarios and success criteria (excluded from analysis):

on the computer screen
with speed and ease-of-use
when I provide a book title
even if the book title is incomplete
so that I can find all books with similar titles

Quality Suggestions

Recommended Verbs

Users Candidates
library user

Objects Candidates
book

Reference
LIB-01

Requirement type
Functional

Found candidate
Users and Objects
Correct and Consistent - Yes
But incomplete!

Data maintenance analysis
Find and fix potential missing and duplicate requirements. Each maintained object of interest usually has one Create, Read, Update and Delete.

<table>
<thead>
<tr>
<th>Object (1) confirm</th>
<th>Create (0)</th>
<th>Read (1)</th>
<th>Update (0)</th>
<th>Delete (0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>book</td>
<td>Missing+</td>
<td>Search library</td>
<td>Missing+</td>
<td>Missing+</td>
</tr>
</tbody>
</table>
BUT WHAT ABOUT AN EPIC?

- Here is an example of an epic from the internet
- How long would it take to make a precise, accurate and repeatable functional size estimate?
- This is how long it took me ...
SCRUM ALLIANCE WEB SITE “PROFILES V1”

- Upload 14 User Stories – 15 minutes
- NLP analysis – 6 seconds
- Summary Result

Analysing the text of software requirements

**PROGRESS**

- Analysing: 14 of 14 (100%)

**FUNCTIONAL SIZE**

- Detected: 30 CFP for 8 sizeable Stories
- Estimated Total: 53 CFP for all 14 Stories

**QUALITY**

- Potential defects found:
  - Ambiguous: 6
  - Missing: 19
  - Duplicated: 1
  - Complex: 0
  - Total: 26

**VALUE OF ANALYSIS**

- Finding defects: $3,900
- Knowing size: $5,250
These estimates use industry benchmark data. We advise that you to use the ScopeMaster derived functional size to **create your own benchmark** data for cost, effort, duration, and quality.

• Need to separate the functional part from the non-functional part – v2.
SCRUM ALLIANCE WEB SITE “PROFILES V2”

Split FP from NFP – 10 minutes
NLP analysis – 6 seconds
Summary Result

Analysing the text of software requirements

- **PROGRESS**
  - Analysing: 14 of 14 (100%)
  - Completed in 6 seconds

- **FUNCTIONAL SIZE**
  - Detected: 27 CFP for 8 sizeable requirements
  - Estimated Total: 47 CFP for all 14 requirements

- **QUALITY**
  - Potential defects found:
    - Ambiguous: 6
    - Missing: 16
    - Duplicated: 1
    - Complex: 0
  - Total: 23

- **VALUE OF ANALYSIS**
  - Finding defects:
    - Value: $3,450
  - Knowing size:
    - Value: $4,725
These estimates use industry benchmark data. We advise that you to use the ScopeMaster derived functional size to create your own benchmark data for cost, effort, duration, and quality.

"PROFILES V1" VERSUS "PROFILES V2"
### Short title

01. Member Profile

### Functional Requirement

As a site member I want to edit my profile

### Notes, scenarios and success criteria (excluded from analysis)

in a semi-structured way so that others can learn about me. That is, I can fill in predefined fields, but also have room for a free-text field or two. (It would be nice to let this free text be HTML or similar.)

### Quality Suggestions

Recommended Verbs
Reword 6 User Stories – 12 minutes
NLP analysis – 6 seconds
Summary Result

Analysing the text of software requirements

**PROGRESS**
Analysing: 14 of 14
100%
6 seconds

**FUNCTIONAL SIZE**
Detected
49 CFP for 14 sizeable requirements
Estimated Total
49 CFP for all 14 requirements

**QUALITY**
Potential defects found:
- Ambiguous: 0
- Missing: 25
- Duplicated: 4
- Complex: 0
Total: 29

**VALUE OF ANALYSIS**
Finding defects:
- $4,350
Knowing size:
- $4,900
“PROFILES V2” VERSUS “PROFILES V3”
<table>
<thead>
<tr>
<th>App</th>
<th>Requirements</th>
<th>Users</th>
<th>Objects</th>
<th>Functional steps</th>
<th>Total CFP</th>
<th>Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCRUMMASTERV1</td>
<td>14</td>
<td>17</td>
<td>46</td>
<td>10</td>
<td>53</td>
<td>59</td>
</tr>
<tr>
<td>SCRUMMASTERV2</td>
<td>14</td>
<td>8</td>
<td>25</td>
<td>9</td>
<td>47</td>
<td>61.4</td>
</tr>
<tr>
<td>SCRUMMASTERV3</td>
<td>14</td>
<td>7</td>
<td>10</td>
<td>15</td>
<td>49</td>
<td>78.4</td>
</tr>
</tbody>
</table>
but ...
### DEFECTS ARE UNMASKED

<table>
<thead>
<tr>
<th>Application</th>
<th>Ambiguous</th>
<th>Missing*</th>
<th>Duplicates*</th>
<th>Defects per Story*</th>
<th>Defects per CFP*</th>
<th>Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCRUMMASTERV1</td>
<td>6 of 14</td>
<td>19</td>
<td>1</td>
<td>1.86</td>
<td>0.49</td>
<td>59</td>
</tr>
<tr>
<td>SCRUMMASTERV2</td>
<td>6 of 14</td>
<td>16</td>
<td>1</td>
<td>1.64</td>
<td>0.49</td>
<td>61.4</td>
</tr>
<tr>
<td>SCRUMMASTERV3</td>
<td>0 of 14</td>
<td>25</td>
<td>4</td>
<td>2.07</td>
<td>0.59</td>
<td>78.4</td>
</tr>
</tbody>
</table>

*All defects are marked as unmasked.*

---

**IWSM Mensura 2019 - Copyright © 2017-2019 Symtech Ltd. All rights reserved. 26/09/2019**
WHAT HAVE I MISSED?

- These may be covered in one of the other epics – but they may not
- There were 5 more pages of User Stories – probably containing 14 US per page – 70 US total?
- So how long to do a full analysis?

<table>
<thead>
<tr>
<th>Object (9) confirm</th>
<th>Create (2)</th>
<th>Read (6)</th>
<th>Update (3)</th>
<th>Delete (0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>application</td>
<td>Duplicate</td>
<td>Missing</td>
<td>Missing</td>
<td>Missing</td>
</tr>
<tr>
<td></td>
<td>02. CSP Application</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>04. CST Application</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>certification status</td>
<td>Missing</td>
<td>06. Article Graphic</td>
<td>Missing</td>
<td>Missing</td>
</tr>
<tr>
<td>class</td>
<td>Missing</td>
<td>07. List Classes SMW</td>
<td>Missing</td>
<td>Missing</td>
</tr>
<tr>
<td>detail</td>
<td>Duplicate</td>
<td>Missing</td>
<td>Missing</td>
<td>Missing</td>
</tr>
<tr>
<td></td>
<td>02. Additional Profile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>05. Additional Train</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>email</td>
<td>Missing</td>
<td>12. Send Email SMWS-</td>
<td>Missing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>01. Email Private SM</td>
<td>Missing</td>
<td></td>
</tr>
<tr>
<td>practicing</td>
<td>Missing</td>
<td>13. Approve applicat</td>
<td>Missing</td>
<td>Missing</td>
</tr>
<tr>
<td>profile</td>
<td>Missing</td>
<td>01. Member Profile SMWS-01</td>
<td>Duplicate</td>
<td>Missing</td>
</tr>
<tr>
<td></td>
<td>As a site member I want to edit my profile...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>01. Member Profile S</td>
<td>10. Profile Private</td>
<td></td>
<td></td>
</tr>
<tr>
<td>site member profile</td>
<td>Missing</td>
<td>Missing</td>
<td>14. Edit member prof</td>
<td>Missing</td>
</tr>
<tr>
<td>training application</td>
<td>Missing</td>
<td>13. Approve applicat</td>
<td>Missing</td>
<td>Missing</td>
</tr>
</tbody>
</table>
• v1 took 15 mins 6 sec
• v2 took 10 mins 6 sec
• v3 took 12 mins 5 sec
• Total 37 mins 17 sec for 14 User Stories
• So maybe 3.5 hours for 70 User Stories
SUMMARY

• 14 Correct and Consistent User Stories in 37 min
• Precise - all User Stories sized ✓
• Accurate - same size for same function ✓
• Repeatable – export / import gives the same result ✓
• But not yet complete.
• About 3.5 hours to estimate all 70
COULD ALEXA HELP?

Computer
start cosmic functional size
Calculate total
Add story
user
manage
my user profile
yes
Calculate total

Add story
Manager
delete
profiles
yes
Calculate total
Goodbye

Thanks to Roland Jacob Patalinjug
CONCLUSIONS

Text and syntax analysis is not enough, **semantic analysis is needed**
Correctness and consistency is not enough, **completeness is needed**
As quality increases **masked defects are exposed**

RECOMMENDATIONS

Committing to a sprint with an incorrect, inconsistent or incomplete set of user stories may be agile but it is certainly not lean
The lean principle of waste reduction needs to be employed
AI techniques can and are available to help.
AI TECHNIQUES FOR SOFTWARE COST ESTIMATION

Dr. Simon Wright

simon@symtech.ltd.uk
QUESTIONS

Why are User Stories returned to the backlog?

Committing to a sprint with an incorrect, inconsistent or incomplete set of user stories may be agile but it is certainly not lean.